

**Schriftliche Stellungnahme des
Verbands deutscher Sprecher:innen e.V. (VDS)
für die Enquetekommission
"Künstliche Intelligenz – Für einen smarten Staat in
der digitalisierten Gesellschaft"
des nordrhein-westfälischen Landtags**

1. Welche Anwendungen generativer KI im Bereich Sprechen sind zurzeit besonders relevant?

- 1.1 Stimmklone
- 1.2 Morphing
- 1.3 Text-to-Speech-Verfahren
- 1.4 Speech-to-Speech-Verfahren
- 1.5 Training
- 1.6 KI ist nicht gleich KI

2. Wie gut sind mittels generativer KI erzeugte synthetische Stimmen?

3. Potenziell faire Rahmenbedingungen für mögliche Einsatzgebiete synthetischer Stimmen auf Basis geltenden Rechts

- 3.1 Persönlichkeitsrecht
- 3.2 Urheberrecht und Leistungsschutzrecht
- 3.3 Das Potenzial synthetischer Stimmen nutzen und dabei geltendes Recht respektieren

4. Aktuelle Probleme und Missstände bei Erzeugung und Anwendung synthetischer Stimmen

- 4.1 Fair Use
- 4.2 Text- und Datamining-Schranke
- 4.3 Eindeutige AGB großer Anbieter:innen
- 4.4 Immer häufiger auftretende Knebelverträge marktführender Auftraggebender
- 4.5 Aussagen wichtiger Marktteilnehmender

5. Welche Auswirkungen hat die Anwendung synthetischer Stimmen bereits heute im Alltag von professionellen Sprecher:innen?

- 5.1 Das „Brot und Butter“-Geschäft wackelt mit Folgen darüber hinaus
- 5.2 Verträge und Machtgefälle
- 5.3 Mangelnde KI-Kompetenz auf allen Ebenen
- 5.4 Rechtsrisiken für Nutzer:innen synthetischer Stimmen und Ansprüche echter Stimmen
- 5.5 KI-Einsatz ohne Mehrwert

6. Blick über den Tellerrand

- 6.1 KI verstärkt potenziell eine rechtliche Zweiklassengesellschaft
- 6.2 Kann KI Kunst?
- 6.3 Auswirkungen auf Arbeitsmarkt und Wirtschaftsleistung
- 6.4 Der Wert geistigen Eigentums ist unschätzbar
- 6.5 Missbrauchspotenzial
- 6.6 Wer profitiert?

6.7 Wie kann KI das Klima schützen?

7. Fazit

1. Welche Anwendungen generativer KI im Bereich Sprechen sind zurzeit besonders relevant?

1.1 Stimmklone

Stimmklone eines Sprechers oder einer Sprecherin werden entweder aus Originalsprachaufnahmen der jeweiligen Person *synthetisiert* oder aus verschiedenen anderen echten Stimmen durch sogenanntes *Morphing*.

1.2 Morphing

Morphing macht es möglich, aus verschiedenen Stimmen eine "neue" *synthetische Stimme* zu erzeugen. Die Originalstimmen, die dazu verwendet wurden, sind nur schwer, meist nicht erkennbar und technisch nicht rückverfolgbar.

1.3 Text-to-Speech-Verfahren

Ein Text wird in ein KI-Modell eingespeist und direkt eine Audiodatei mit einer *synthetischen Stimme* generiert, es bedarf einer aufwändigen Postproduktion, um einen akzeptablen Standard zu erreichen. Macht die Produktion nicht günstiger bei wesentlich schlechterer Qualität im Vergleich zu einer echten Stimme.

1.4 Speech-to-Speech-Verfahren

Synthetisierte *Stimmmasken* werden auf echte Sprecher:innen gesetzt. Findet bereits Anwendung, wenn z.B. Jürgen Klopp im Rahmen einer Werbung in mehreren Sprachen mit seiner Stimmfarbe sprechen soll. Dann lässt man verschiedene Nativespeaker den Text einsprechen und Jürgen Klopps synthetisierte *Stimmmaske* wird darübergezogen. Klingt nicht so authentisch wie versprochen, sorgt aber für Faszination und weckt bei manchen Synchronfirmen die Ambition, z.B. Brad Pitt mit seiner Stimmfarbe Deutsch sprechen zu lassen. Dieses Versprechen kann allerdings nicht gehalten werden, da jede Sprache verschiedene Anforderungen an den Sprechapparat stellt und somit auch die Stimme einer Person anders klingt, je nachdem welche Sprache sie spricht. Zudem besteht eine Stimme aus weit mehr als ihrer Farbe / ihrem Timbre. Ebenso ausschlaggebend für eine Stimmidentität sind z.B. Stimmführung, Modulation, Sprachmelodie, Timing, Atmung, Druck, Lebenserfahrung, Persönlichkeit, Charme, Humor uvm.

1.5 Training

Die Basis all dieser Anwendungen ist das umfangreiche *Training generativer KI-Modelle*, sogenannter *Large Language Models*, kurz *LLMs*. Auf Basis enormer Datenmengen berechnet ein Modell auf Fragen und Anweisungen (*Prompts*) die wahrscheinlichsten Antworten und Angebote, im Falle von *synthetischen Stimmen* also z.B. Tonfall, Sprechtempo, Pausen, Betonungen etc.

1.6 KI ist nicht gleich KI

Wir möchten am Anfang dieses Berichts betonen, dass wir unser Augenmerk nicht auf den Einsatz *analytischer*, sondern *generativer KI-Modelle* legen, die damit verbundenen Chancen, Risiken und Bedenken aus Sicht unseres Verbandes.

2. Wie gut sind mittels generativer KI erzeugte synthetische Stimmen?

Noch kommen *synthetische Stimmen* in allen unter Punkt 1 beschriebenen Anwendungsformen nicht an die Qualität echter Stimmen heran und es wird sehr wahrscheinlich auch so bleiben.

Wir als Verband mit unserem Wissen um die Komplexität des Sprechens und kreativer Arbeit im Allgemeinen, sind der festen Überzeugung: nur Menschen sind in der Lage, Gefühle in Worte zu fassen und eine empathische Verbindung mit dem Publikum und den Zuhörer:innen herzustellen, Texte mit mehreren Bedeutungsebenen originell und mit Individualität zu interpretieren. KI kann die Nuancen, die für eine authentische Ausführung dieser Aufgabe in hoher Qualität notwendig sind, weder meistern noch verstehen.

Generative KI-Modelle arbeiten wie unter 1.5 beschrieben mit Wahrscheinlichkeiten auf Basis ihrer Trainingsdaten, d.h. dass z.B. eine *synthetische Stimme*, die einen Text hörbar macht, dies in allen Aspekten auf die naheliegendste, in den Trainingsdaten am häufigsten vorkommende Weise tun wird. Das ist auch der Grund, weshalb eine *synthetische Stimme* im Ergebnis im wahrsten Sinne des Wortes mittelmäßig klingt. Denn das wahrscheinlichste ist im kreativen Arbeiten meist auch das langweiligste.

Hohe Kreativität zeichnet aus, dass jede kleinste Entscheidung höchst individuell, originell und authentisch getroffen wird. Das können nur Menschen bieten - mit ihrer eigenen Lebensgeschichte, ihren spezifischen Erfahrungen, ihrer Berufsausbildung etc.

Es ist aber nicht zu leugnen, dass durch das Einspeisen enormer Datenmengen, in unserem Fall Stimmproben aus verschiedensten Quellen, die KI-Modelle definitiv besser werden bzw. natürlicher klingen. Man spricht hier vom sogenannten *Neural Learning*. Was genau in einem *künstlichen neuronalen Netz* vor sich geht, ist technisch nicht einseh- geschweige denn nachvollziehbar: Die Lernerfolge, die es durch verschiedene Trainingsdaten erzielt, sind daher nicht rückgängig zu machen, auch wenn die ursprünglichen Trainingsdaten gelöscht werden. Die Modelle memorisieren Daten und es gibt keine Möglichkeit der Entfernung¹.

¹ siehe Tim W. Dornis, Sebastian Stober: *Urheberrecht und Training generativer KI-Modelle. Technologische und juristische Grundlagen*. Baden-Baden: Nomos 2024, online unter: <https://www.nomos-elibrary.de/de/10.5771/9783748949558/urheberrecht-und-training-generativer-ki-modelle?page=1>

3. Potenziell faire Rahmenbedingungen für mögliche Einsatzgebiete synthetischer Stimmen auf Basis geltenden Rechts

3.1 Persönlichkeitsrecht

Die Stimme ist eine biometrische Eigenschaft eines Menschen und daher vom Persönlichkeitsrecht geschützt, das gilt für professionelle Sprecher:innen wie auch für alle anderen Menschen.

3.2 Urheberrecht und Leistungsschutzrecht

Die Darbietungen professioneller Sprecher:innen sind zudem vom Urheberrecht geschützt, da sie in den meisten Fällen eine künstlerische Leistung darstellen. Sollten sie nicht urheberrechtlich geschützt sein, so sind sie mindestens durch das Leistungsschutzrecht geschützt.

3.3 Das Potenzial synthetischer Stimmen nutzen und dabei geltendes Recht respektieren

Der Verband Deutscher Sprecher:innen e.V. (VDS) steht *synthetischen Stimmen* und KI allgemein nicht generell ablehnend gegenüber, im Gegenteil: z.B. in der Barrierefreiheit beim Hörbarmachen von Zeitungsartikeln oder anderweitiger Informationsvermittlung für Menschen mit Sehbehinderung, häufigen Aktualisierungen in E-Learnings oder Supermarktwerbung, bei Krankheit der Sprecher:innen, Produktionsprozessen mit hohem Zeitdruck etc. gibt es durchaus potenzielle Anwendungsfelder.

Der Verband hat hierfür einen KI-Gagenkompass² entwickelt, in dem er Standards und Rahmenbedingungen für einen gesetzeskonformen, fairen, sicheren und sinnvollen Einsatz *synthetischer Stimmen* formuliert.

So muss u.a. gewährleistet sein, dass Sprechende KI-Anwendungen mit ihren Stimmen explizit zustimmen müssen und Nutzungsumfang und -zeitraum klar definiert, begrenzt und angemessen vergütet werden. Deutsche Datenschutzstandards müssen respektiert werden und es gilt generell absolute Transparenz über die Nutzung der echten Stimme im Rahmen künstlicher Intelligenz.

Davon, seine Stimme unkontrollierbar zu KI-Trainingszwecken zur Verfügung zu stellen (siehe 2. unten), rät der Verband ab bzw. empfiehlt seinen Mitgliedern zumindest eine massive Vergütung, da man seine Stimme mit all ihren individuellen Charakteristika, in unserem Fall die Existenzgrundlage, potenziell aufgibt.

Um eventuellem Missbrauch vorzubeugen, aber auch Auftraggebende für das Thema zu sensibilisieren und Transparenz zu schaffen, hat unser Verband zusammen mit seinem internationalen Dachverband UNITED VOICE ARTISTS (UVA) eine Vertragsklausel entwickelt, die sämtliche *generative KI-Anwendungen* von Sprachaufnahmen ohne die Zustimmung der Sprecher:innen ausschließt³. Ob diese seitens der Auftraggebenden

² <https://www.sprecherverband.de/sprachsynthese-generative-ki/>

³ <https://www.sprecherverband.de/ki-ausschluss-rechteuebertragungen/>

akzeptiert wird oder nicht, ist ein guter Indikator für deren Vertrauenswürdigkeit in puncto *synthetische Stimmen*.

4. Aktuelle Probleme und Missstände bei Erzeugung und Anwendung synthetischer Stimmen

Leider finden die unter 3.3 angerissenen Prinzipien und Vertragsinhalte in der Praxis bislang kaum bzw. zu selten Anwendung: Die meisten Anbieter:innen synthetischer Sprachmodelle sitzen in den USA oder Asien und haben ein völlig anderes Verständnis von Datenschutz und Urheberrecht, d.h. es besteht wenig Respekt vor dem geltenden deutschen Recht. Große Auftraggebende verweigern meist den Ausschluss *generativer KI-Anwendungen* in ihren Verträgen und führende Anbieter:innen *synthetischer Stimmen* geben keine oder kaum Informationen über die Provenienz ihrer Trainingsdaten heraus. In unzähligen Fällen werden Mediatheken, Soziale Netzwerke, Websites von Sprecher:innen etc. *gecrawlt* oder ausgesaugt, um mit den hochwertigen Stimmdateien LLMs zu *trainieren*. Sogar, wenn diese per *robots.txt* blockiert wird⁴. Aus unter 2. genannten technischen Gründen ist dieses unrechtmäßige *Crawling* allerdings sehr schwer nachweisbar, jedoch sprechen viele Indizien dafür:

4.1. Fair Use

In den USA wird die *Fair Use-Doktrin* häufig so verstanden, dass öffentlich zugängliche Inhalte im Internet auch für das Training von KI-Systemen genutzt werden dürfen. Diese Praxis wird von vielen Unternehmen tatsächlich so umgesetzt und mit dem weiten Anwendungsbereich von *Fair Use* begründet. Allerdings ist diese Auslegung rechtlich nicht eindeutig abgesichert, sondern umstritten und wird von unterschiedlichen Akteuren – darunter Rechteinhaber und Juristen – teils deutlich angefochten. Nach deutschem Urheberrecht handelt es sich potenziell um eine Persönlichkeits- und/oder Urheberrechtsverletzung, denn nur weil Sprecher:innen z.B. mit Hörproben auf ihrer Website ihr Können bewerben, ist damit noch lange kein Einverständnis für andere Nutzungsformen gegeben. Der Inhalt einer Mediathek wiederum ist auf einen bestimmten Zeitraum und ein bestimmtes Publikum (Beitragszahler:innen) beschränkt.

4.2 Text- und Datamining-Schranke

Die *Text- und Datamining-Schranke* des Urheberrecht erlaubt die Erhebung und Verarbeitung großer Datenmengen zu nicht-kommerziellen Forschungszwecken. Da die Trainingserfolge technisch so schwer bzw. nicht zurückzuverfolgen sind, nicht zurückgenommen werden können und damit auch der Profit, den sie bei

⁴ <https://www.plagiarismtoday.com/2025/07/23/chatgpt-ignores-robots-txt-rehashes-my-column/>

Kommerzialisierung generieren, nicht nachzuverfolgen ist, wird der Forschungszweck als alleinige Motivation vorgeschoben und damit versucht, die unrechtmäßige Nutzung der Daten zu legitimieren.

4.3 Eindeutige AGB großer Anbieter

In unübersichtlich langen AGB führender Anbieter:innen *synthetischer Stimmen* wie z.B. ElevenLabs steht eindeutig, dass hochgeladenes Material zum *Training* verwendet werden kann, auch wenn dies mit irreführenden Sicherheitsversprechen verschleiert wird. Aber auch Soziale Netzwerke wie Meta und sogar Dienstleister:innen völlig anderer Services wie z.B. WeTransfer oder eBay haben schwer auffindbare, extrem umfangreiche Zustimmungen zu KI-Trainingszwecken in ihren AGB formuliert. Auch wenn es Möglichkeiten zum Widerspruch gibt, sind diese unserer Meinung nach unzureichend, denn meistens hat das Training vor dem Widerspruch bereits stattgefunden und ist wie unter Punkt 2 nicht rückgängig zu machen. Wir fordern die Verpflichtung der Anbieter:innen, eine explizite Einwilligung bei Rechteinhaber:innen einholen zu müssen.

4.4 Immer häufiger auftretende Knebelverträge marktführender Auftraggebender

Immer mehr Vertragswerke, die uns in unserem beruflichen Alltag begegnen, beinhalten explizite oder zumindest implizite Zustimmungen zu verschiedenen KI-Anwendungen mit unserer Stimme, sehr häufig *Training*. Der Fakt, dass große Auftraggebende (z.B. Streamingdienste, Tech-Unternehmen, Videospieldpublisher, Produktionsfirmen) auf diese Vertragsinhalte beharren und unsere eigens zu diesem Zweck entwickelte KI-Ausschlussklausel meist ablehnen, zeigt das große Interesse an hochwertigen Trainingsdaten und lässt vermuten, dass damit potenzielle *generative KI-Anwendungen* gefüttert werden sollen.

4.5 Aussagen wichtiger Marktteilnehmender

Aussagen marktrelevanter Branchengrößen in Verhandlungen mit unserem Dachverband UNITED VOICE ARTISTS bestätigen die unter 4.4 beschriebene Vermutung. Uns ist sogar schon der Fall begegnet, bei dem eine große deutsche Filmproduktion die Geschäftsidee geäußert hat, für sämtliche bisher produzierten Werke *KI-Trainingsrechte* zu verkaufen, ohne jegliches Bewusstsein für juristische und monetäre Ansprüche auf Seiten der Kreativen, mit deren Leistung, Erscheinungsbild und Stimme diese Werke erstellt wurden.

5. Welche Auswirkungen hat die Anwendung synthetischer Stimmen bereits heute im Alltag von professionellen Sprecher:innen?

5.1 Das "Brot und Butter"-Geschäft wackelt mit Folgen darüber hinaus

Im Low-Budget-Bereich werden immer häufiger synthetische Stimmen eingesetzt, z.B. bei Bauanleitungen auf YouTube, vorgeschalteter Onlinewerbung vornehmlich kleiner Unternehmen etc. und überall da, wo Endkund:innen keinen Qualitätsanspruch haben bzw. Rezipient:innen mangelnder Qualitätsanspruch unterstellt wird.

Dies verschlechtert die allgemeine Auftragslage für viele Sprecher:innen im sogenannten "Brot und Butter"-Geschäft. Man könnte argumentieren, dass dieser Umstand bei den Sprechenden Freiräume für kreativ anspruchsvollere Projekte schafft. Da die Sprecher:innen an den Erträgen *synthetischer Stimmen* aber nicht beteiligt werden, obwohl sie sehr wahrscheinlich zu deren Entstehung und Verbesserung beigetragen haben (siehe Punkt 4) und solche Aufträge vielen überhaupt erst die finanzielle Freiheit geben, auch kreativ anspruchsvollere, oft schlechter bezahlte Projekte zu realisieren, ist das Gegenteil der Fall. Wir haben hier also einen negativen Dominoeffekt von den einfachen in die anspruchsvollen, in Zukunft weiterhin gefragten Sprechdienstleistungen (Radio, Fernsehen, Kulturveranstaltungen, Premiumwerbung etc.) mit der Folge, dass bei weiterhin unreguliertem Einsatz *synthetischer Stimmen* im Low-Budget-Bereich auch dort die Expertise künftig fehlen könnte - weil die verbliebenen anspruchsvollen Aufträge wirtschaftlich auch für sehr gute Sprecher:innen nicht mehr tragbar werden und auch die Möglichkeit der Weiterbildung der Sprechenden durch Erfahrung bei kleinen Aufträgen wegfällt.

5.2 Verträge und Machtgefälle

Die unter 4.4 bereits erwähnten Verträge mit fragwürdigen KI-Klauseln führen dazu, dass viele Sprecher:innen aus Protest auf Jobs verzichten, da Vertragsänderungen in den allermeisten Fällen seitens der Auftraggeber:innen verweigert werden. Gerade im Bereich der Videospielvertonung (ein großer Markt mit hohem Bedarf an vielen verschiedenen Stimmen) und auch bei einigen größeren Synchronfirmen werden das Machtgefälle und die ohnehin schon prekäre Arbeitssituation vieler Sprecher:innen ausgenutzt, um diese doch zur Unterschrift zu bewegen: Es wird etablierten Sprecher:innen mit Umbesetzung gedroht, Newcomer:innen werden nicht ausreichend über die Verträge aufgeklärt, *non-disclosure-agreements* vorgeschaltet, um den Austausch in der Branche darüber zu unterbinden, Falschinformationen werden verbreitet, unsere Verbandsarbeit diffamiert, verschiedene Standorte versucht, gegeneinander auszuspielen, in dem man z.B. als Kölner Studio nur Berliner Sprecher:innen kontaktiert.

Wir als Verband versuchen diesem Gebaren mit viel Aufklärungsarbeit, Vernetzung, Aufrufen zu Weitsicht und Solidarität und verstärkter Öffentlichkeitsarbeit

entgegenzuwirken. Zuletzt mit einer Petition, die für viel Aufmerksamkeit auch über unser Branche hinaus gesorgt hat⁵.

Auch wenn wir damit bereits Erfolge erzielen konnten und unsere Ausschlussklausel für *generative KI* immer öfter in Verträge aufgenommen wird, kann nur jede einzelne Person ihre Ideale verteidigen. Ein Streik ist so gut wie unmöglich, da es sich bei den meisten Aufträgen nicht um Anstellungsverhältnisse handelt und es nur selten so etwas wie Tarifverträge gibt.

Alles in allem kann man beim Thema Verträge zusammenfassen: Sprecher:innen verlieren ihre Jobs nicht, weil die KI es besser kann, sondern weil sie sich weigern, die KI zum Nulltarif dazu zu befähigen und große Auftraggeber:innen ihre Machtposition ausnutzen, um tief ins Persönlichkeits- und Urheberrecht der Sprechenden einzugreifen.

5.3 Mangelnde KI-Kompetenz auf allen Ebenen

Die KI-Kompetenz, die der *AI-ACT* Unternehmen und ihren Angestellten vorschreibt, ist in den meisten Fällen schlichtweg nicht vorhanden (siehe auch 4.5). Immer wieder begegnen wir unbedachter und ungefragter Nutzung von KI-Anwendungen, bei denen mit den Stimmdateien der Sprecher:innen *trainiert* wird - durch Studios, Produktionsfirmen, durch Rundfunkanstalten oder auch Synchronfans, die gerne mit der Stimmmaske ihres Idols sprechen möchten.

Die Schäden, die dadurch entstehen, sind enorm, denn wie unter 1.5. bereits erklärt ist die Basis sämtlicher Anwendungen von *synthetischen Stimmen* und des damit erwirtschafteten Profits *Training*.

Diese mangelnde Kompetenz findet sich leider auch im Einsetzungsbeschluss dieser Enquetekommission wieder, der viele Marketing-Versprechen aus der Industrie ohne Belege oder objektive Einordnung wiederholt und als Fakten darstellt⁶.

5.4 Rechtsrisiken für Nutzer:innen synthetischer Stimmen und Ansprüche echter Stimmen

Wir verstehen, dass *synthetische Stimmen* für Unternehmen und Einzelpersonen mit geringem Budget, die sich keine professionellen Sprechenden leisten können, eine scheinbar günstige Alternative z.B. für Eigenwerbung oder im Kund:innendialog darstellen. Diese Anwendungen bergen allerdings das Risiko für Nutzer:innen, gegen Persönlichkeits- oder Urheberrecht zu verstoßen. Solche Verstöße finden statt, sobald eine generierte *synthetische Stimme* einer echten Stimme gleicht oder auch nur in Teilen ähnelt - selbst wenn dieser Stimmklon mittels Morphing anderer Stimmen produziert wurde! Es muss nicht einmal die Originalstimme eingespeist worden sein, sobald der *Output* einer echten Stimme ähnelt ohne dass der/die Rechteinhaber:in dieser echten Stimme gefragt

⁵ <https://www.openpetition.de/petition/online/schuetzt-die-kunst-vor-ki-deinestimmefuerechtestimmen>

⁶ siehe <https://dcthemedian.substack.com/p/five-myths-about-genai-in-the-enterprise>

wurde, machen sie sich strafbar. Auch hier laufen bereits seit Jahren mehrere Klagen und der Synchronstimme von Bruce Willis wurde vor kurzem Recht zugesprochen⁷.

Als Verband möchten wir in Zusammenarbeit mit Forschungsinstituten und Unternehmen technische Möglichkeiten entwickeln, Stimmklone und *Deepfakes* aller Art automatisiert aufzuspüren und strafrechtlich zu verfolgen. Da es sich hierbei um die Durchsetzung und Wahrung allgemein geltenden Persönlichkeitsrechts und Sicherheitsfragen handelt, betrifft diese Aufgabe in hohem Maße die gesamte Gesellschaft und ist von zentraler Bedeutung für eine große Zahl von Bürger:innen. Vor diesem Hintergrund wäre eine finanzielle Unterstützung bzw. Förderung seitens des Gesetzgebers nicht nur wünschenswert, sondern auch ein klares Signal zur Stärkung von Rechtsstaatlichkeit und Sicherung des Gemeinwohls.

5.5 KI-Einsatz ohne Mehrwert

Auch wenn sich die Strafverfolgung bei unbefugter Nutzung echter Stimmen zurzeit noch sehr schwierig gestaltet, beobachten wir in der freien Wirtschaft bereits eine Rückbesinnung auf echte Stimmen bei Kund:innen mit einem Mindestmaß an Qualitätsanspruch. Grund dafür ist, dass Vertonungen mit *synthetischen Stimmen* einen aufwändigen technischen Postproduktionsprozess durchlaufen müssen, der eventuelle Ersparnisse beim Sprecher:innenbudget wieder verschlingt. Der Produktionsprozess wird also nicht effektiver, er wird länger und teurer, die Kosten von einer kreativen zu einer technischen Leistung verschoben. Das Ergebnis wird weniger interessant, weil es nicht intuitiv entsteht, sondern generiert ist.

Immer häufiger begegnen wir dem kritiklosen Einsatz *generativer KI* in verschiedenen Produktionsschritten aber bei öffentlich-rechtlichen Rundfunkanstalten, z.B. bei Übersetzungen für zu vertonende Texte. Auch wenn diese nachträglich redaktionell geprüft werden, kommen diese Übersetzungen nicht einmal annähernd an menschengemachte Übersetzungen heran. Vielleicht lässt sich dadurch Geld sparen, aber zum Preis immer schlechter werdender und fehlerhafter Inhalte. Aufgrund des Bildungsauftrags der Sender und der Finanzierung mithilfe der Bürger:innen ist dies aus unserer Sicht höchst problematisch.

⁷ siehe <https://ra-juedemann.de/auch-die-synchronstimme-ist-durch-das-allg-persoenlichkeitsrecht-geschuetzt-lg-berlin-urteil-vom-20-august-2025-ii-2-o-202-249/>

6. Blick über den Tellerrand

6.1 KI verstärkt potenziell eine rechtliche Zweiklassengesellschaft

Wir als Verband beobachten die enorme Einflussnahme des Gesetzgebers auf allen Entscheidungsebenen durch führende KI-Unternehmen in Form von großangelegter Lobbyarbeit. In Großbritannien konnte man jüngst gut beobachten, wie die KI-Industrie versucht, z.B. das Urheberrecht in ihrem Sinne immer weiter aufzuweichen⁸.

Auch führen aggressives Marketing samt haltloser Versprechen zu einer gewissen Panik bei Politik wie auch Unternehmen, "abgehängt" zu werden. Es gibt unzählige sinnvolle KI-Anwendungen gerade im Bereich *analytischer KI*, die Arbeit vollführt, die Menschen ob der schiereren Datenmengen niemals leisten könnten. Es kann aber nicht sein, dass Persönlichkeits- oder Urheberrecht zugunsten großer Unternehmen gebeugt werden. Beide sind ein Grundpfeiler unserer Demokratie und müssen für einzelne Personen wie Unternehmen gleichermaßen gelten.

6.2. Kann KI Kunst?

Nein. Künstler:innen reflektieren und kommentieren seit jeher die Gesellschaft, in der sie leben, schaffen Empathie und Aufmerksamkeit. Sie sind somit integraler Bestandteil einer freiheitlich-demokratischen Gesellschaft. Der Einsatz von künstlich erschaffenem Text-, Bild- und Tonmaterial (*Deepfakes*) kann diese Funktion nicht ausfüllen, da lediglich bereits bestehende Werke maschinell wiedergekaut werden. Eine künstlerische Intention kann man einer KI nicht einprogrammieren, sie soll gefallen, sich verkaufen. Aber Kunst darf nicht nur gefallen wollen! Sie muss auch provozieren, Fragen aufwerfen, Debatten entfachen. Das kann sie nur, wenn sie menschengemacht ist.

6.3 Auswirkungen auf Arbeitsmarkt und Wirtschaftsleistung

Speziell in der Kultur- und Medienlandschaft führt der unregulierte Einsatz *generativer KI* nicht nur zu den bereits beschriebenen Qualitätsverlusten, sondern bedroht Existenzgrundlagen einer großen Menge kreativ arbeitender Menschen, da bei den meisten ein Mix aus künstlerisch anspruchsvollen und weniger anspruchsvollen Projekten die Arbeit erst wirtschaftlich tragbar macht.

Werden künstlerisch weniger anspruchsvolle Arbeiten durch *generative KI-Anwendungen* übernommen, die wahrscheinlich mit dem Können ebenjener Kreativen *trainiert* wurden ohne dass diese dafür angemessen vergütet wurden, wird eine große Umorientierung stattfinden mit fatalen Folgen für die Bereiche, in denen menschliche künstlerische Expertise unabdingbar bleiben wird - eine ernsthafte wirtschaftliche Bedrohung gerade auch für den Medienstandort NRW.

⁸ siehe <https://www.heise.de/news/Urheberrecht-Grossbritannien-will-geschuetzte-Werke-fuers-KI-Training-freigeben-10217463.html>

6.4 Der Wert geistigen Eigentums ist unschätzbar

Die KI-Industrie hat dies längst begriffen und speist es ungezügelt in ihre Systeme. Auch die Politik sollte ihr Bewusstsein dafür schärfen, welch unschätzbaren Wert die Produktion und der Schutz geistigen Eigentums für eine Gesellschaft und auch die Wirtschaft hat. Missachtung und Entwertung geistigen Eigentums, wie sie zurzeit stattfinden, lassen den Eindruck entstehen, dass es sich nicht mehr lohnt, geistiges Eigentum zu produzieren. Das wäre ein gefährliches Signal mit potenziell nicht zu ersetzenden Verlusten, nicht nur in der wirtschaftlich so relevanten Kreativbranche NRWs, sondern auch in seiner Forschung, Innovation und Publizistik.

6.5 Missbrauchspotenzial

Deepfakes allgemein, aber gerade auch mit *synthetischen Stimmen* haben ein hohes Missbrauchspotenzial und stellen somit ein Sicherheitsrisiko nach Innen wie Außen dar. Die Möglichkeiten reichen von krimineller Nutzung bis hin zu gezielter politisch motivierter Desinformation. Forschung, Informationsfluss, Publizismus, Journalismus und auch Übersetzungen müssen daher, ebenso wie jegliche Kunstform, unter menschlicher Kontrolle bleiben.

6.6 Wer profitiert?

Aktuell konzentrieren sich die wirtschaftlichen Vorteile *generativer KI-Technologien* hauptsächlich auf Unternehmen in den USA und Asien, während geistiges Eigentum aus Europa häufig als Trainingsgrundlage für diese Systeme genutzt wird. Die Profiteure sitzen also meist in Ländern, in denen Demokratie auch durch den missbräuchlichen Einsatz von KI entweder konkret gefährdet ist oder noch nie existiert hat. Dieser umfassende intellektuelle Raubbau kann weder im Sinne unserer freiheitlich-demokratischen Gesellschaft noch eines wirtschaftsstarken und sicheren Europas sein und ihm muss schnellstens entgegengesteuert werden.

Europa, Deutschland und NRW sollten den Anspruch haben, zu zeigen, wie es besser geht: KI-Modelle fördern, die legal, demokratisch und fair sind - mit echten Vorteilen für eine Mehrheit der Gesellschaft, nicht nur für Großunternehmen und deren Besitzende. Dazu zählt auch die Priorisierung von Grundbedürfnissen wie Wasser, Energie und Raum für Bürger:innen gegenüber Unternehmensinteressen⁹.

6.7 Wie kann KI das Klima schützen?

Die Folgen des unregulierten Einsatzes gerade *generativer KI-Modelle* für das Klima sind bereits jetzt besorgniserregend und es scheint uns daher unumgänglich, sehr klar zu

⁹ siehe bspw. <https://www.n-tv.de/wirtschaft/Rechenzentren-wuergen-Londoner-Wohnungsbau-ab-article23505495.html>

definieren, wo der Einsatz von KI sinnvoll ist und wo er eine Verschwendung, ja Zerstörung von Ressourcen darstellt¹⁰.

Ressourceneffizienz und tatsächlicher Mehrwert sollten wesentliche Kriterien bei der Förderung der KI-Industrie sein. Nur dann kann KI vielleicht ihren Teil dazu beitragen, das Klima zu schützen. Hierbei sehen wir die Potenziale überwiegend im Feld *analytischer KI*.

7. Fazit

Synthetischen Stimmen in jeglicher Form kommen qualitativ nicht an echte Stimmen heran. In gewissen Anwendungsgebieten wie z.B. der Barrierefreiheit und unter fairen und transparenten Rahmenbedingungen, können sie für Auftraggeber:innen wie Sprecher:innen eine nützliche Ergänzung zu echten Stimmen sein. Bei niedrigem Qualitätsanspruch seitens der Auftraggeber:innen und Rezipient:innen und Missachtung geltenden Rechts können Sie aber ein potenzieller Jobkiller für professionelle Sprecher:innen mit gefährlichen Dominoeffekten in alle Bereiche des professionellen Sprechens sein.

Besonders kritisch zu betrachten ist die bisherige Erzeugung *synthetischer Stimmen* durch illegales *Training* auf Basis umfangreichen Datenklau und Urheberrechtsverletzungen. Anbietende *synthetischer Stimmen* müssen ihre Trainingsdaten offenlegen und im Falle von Rechtsbrüchen Kompensationen an die Geschädigten zahlen. Ein Reset unter fairen Rahmenbedingungen ist dringend vonnöten. Diese müssen beinhalten, dass die Einwilligung zum Training *generativer KI-Modelle* von Rechteinhaber:innen explizit eingeholt und bei Zustimmung angemessen vergütet werden muss. Es ist unzumut- und schlichtweg nicht realisierbar, dass einzelne Personen alle Unternehmen dieser Welt, die Ambitionen im Bereich *generativer KI* haben, recherchieren, kontaktieren und ihnen die Nutzung ihrer Daten auf eigene Initiative untersagen müssen. Auch dürfen Arbeitsverhältnisse zwischen großen Unternehmen und einzelnen Sprecher:innen nicht dazu missbraucht werden, letztere zur Aufgabe ihrer Persönlichkeits- und Urheberrechte zu drängen.

KI-Kompetenz in Unternehmen, Institutionen oder auch öffentlich-rechtlichen Rundfunkanstalten sollte nicht nur mit Blick auf einen produktivitätssteigernden Umgang gefördert werden, sondern auch auf Rechtskonformität. Der Enquetekommissionseinsetzungsbeschluss der Fraktionen von CDU und FDP wie auch

¹⁰ siehe <https://netzpolitik.org/2025/generative-ki-klimaschaedlich-by-design/> oder auch <https://www.boell.de/de/2025/01/08/digitalisierung-kuenstliche-intelligenz-und-wasserverschwendung> sowie <https://www.greenpeace.de/ueber-uns/loesungen-finden/kuenstliche-intelligenz-energieverbrauch-und-umweltauswirkungen>

der damit verbundene Fragenkatalog für die Sachverständigen weisen hier in unseren Augen eine doch recht unausgewogene Gewichtung auf.

Wir empfehlen der Landesregierung, eine KI-Industrie zu fördern, die, im Gegensatz zu der in den USA und Asien, geltendes Recht respektiert und schützt, die sich auf sinnvolle Einsatzgebiete mit echtem Mehrwert für Mensch, Gesellschaft, Wirtschaft und Sicherheit fokussiert und begrenzte ökologische Ressourcen klug einsetzt.

Das wären Alleinstellungsmerkmale, mit denen man sich von der ausländischen Konkurrenz absetzen und mit der man zudem den zu Recht großen skeptischen Teil der Bevölkerung überzeugen könnte. Dafür müsste letztlich folgendes klar belegt werden: Der Nutzen für die Gesellschaft ist weit größer als die Verluste. Dies kann bisher nicht nachgewiesen werden.

Wir bedanken uns dafür, dass wir unsere Expertise in die Kommission einbringen durften und stehen für weiteren Austausch gerne zur Verfügung.

Köln, den 3. September 2025



Verband Deutscher Sprecher:innen e.V.
Anna-Sophia Lumpe (1. Vorsitzende)
Weinsbergstr. 189, 50823 Köln

<https://www.sprecherverband.de>
vds-vorstand@sprecherverband.de

Amtsgericht Krefeld | Register-Nr.: VR 4853
Vorstand: Anna-Sophia Lumpe, David M. Schulze, Cédric Cavatore, Miriam Kaufmann